Pósa's Conjecture for graphs of order at least 2×10^8

Phong Châu * Louis DeBiasio † H.A. Kierstead ‡

January 20, 2013

Abstract

In 1962 Pósa conjectured that every graph G on n vertices with minimum degree $\delta(G) \geq \frac{2}{3}n$ contains the square of a hamiltonian cycle. In 1996 Fan and Kierstead proved the path version of Pósa's Conjecture. They also proved that it would suffice to show that G contains the square of a cycle of length greater than $\frac{2}{3}n$. Still in 1996, Komlós, Sárközy, and Szemerédi proved Pósa's Conjecture, using the Regularity and Blow-up Lemmas, for graphs of order $n \geq n_0$, where n_0 is a very large constant. Here we show without using these lemmas that $n_0 := 2 \times 10^8$ is sufficient. We are motivated by the recent work of Levitt, Szemerédi and Sárközy, but our methods are based on techniques that were available in the 90's.

1 Introduction

The *square* H^2 of a graph H is obtained by joining all pairs $\{x, y\} \subset V(H)$ with distance $dist(x, y) = 2$ in H. If H is a path (cycle) then H^2 is called a *square* path (cycle). Now fix a graph $G = (V, E)$ on n vertices. We say that $v_1 \dots v_t$ is a *square* path (cycle) in G if $v_1 \dots v_t$ is a path (cycle) in G and its square is contained in G. In 1962 Pósa [5] conjectured:

Conjecture 1. *Every graph G with $\delta(G) \geq \frac{2}{3}|G|$ contains a hamiltonian square cycle.*

During the 90's there were numerous partial results on Pósa's conjecture. Here we review a number that have a direct impact on this paper. Fan and Kierstead [6, 7, 8] proved the following three theorems. The first is a connecting lemma that immediately yields an approximate version of Pósa's conjecture.

*Department of Mathematics and Computer Science, Glendale Community College, Glendale AZ 85302, USA. e-mail: phong.chau@gcmail.maricopa.edu

†School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287, USA. e-mail address: louis@mathpost.asu.edu

‡School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85287, USA. e-mail address: kierstead@asu.edu; research supported in part by NSF grant DMS-0901520.

Theorem 2 (Fan and Kierstead [6]). *For every $\epsilon > 0$ there exists a constant m such that for every graph G with $\delta(G) \geq (\frac{2}{3} + \epsilon)|G| + m$ and every pair e_1, e_2 of disjoint ordered edges, G has a hamiltonian square path starting with e_1 and ending with e_2. In particular, G has a hamiltonian square cycle.*

We shall need two ideas from this paper—*weak reservoirs* [1], and *optimal* square paths and cycles—which will be presented in the next section. Roughly, given a graph G on n vertices, a weak reservoir is a small fraction R of the vertex set $V(G)$ such that $|N \cap R| \approx |N||R|/n$ for any neighborhood $N := N(v)$. Weak reservoirs were used to connect long square paths contained in $V(G) \setminus R$. The second theorem is a path version of Pósa's Conjecture.

Theorem 3 (Fan and Kierstead [7]). *Every graph G with $\delta(G) \geq \frac{2|G|-1}{3}$ contains a hamiltonian square path.*

The third theorem shows that $V(G)$ can be partitioned into at most two square cycles.

Theorem 4 (Fan and Kierstead [8]). *Suppose G is a graph with $\delta(G) \geq \frac{2}{3}|G|$. If G has a square cycle of length greater than $\frac{2}{3}|G|$ then G has a hamiltonian square cycle. Moreover, $V(G)$ can be partitioned into at most two square cycles, each of length at least $\frac{1}{3}|G|$.*

The proofs of Theorems 3 and 4 are based on optimal paths and cycles, but do not use weak reservoirs. Theorem 4 is essential to this paper, because it allows our constructions to terminate as soon as we get a square cycle of length greater than $\frac{2}{3}|G|$.

Next came a major breakthrough. Komlós, Sárközy and Szemerédi proved their famous Blow-up Lemma [13], and used it and the Regularity Lemma [19] to prove:

Theorem 5 (Komlós, Sárközy and Szemerédi [12]). *There exists a constant n_0 such that every graph G with $|G| \geq n_0$ and $\delta(G) \geq \frac{2}{3}|G|$ has a hamiltonian square cycle.*

Their proof has the following structure. First they determine extremal configurations that are very close to being counterexamples, but because of the tightness of the degree condition, cannot achieve this status. (For example, if the independence number $\alpha(G) > \frac{1}{3}|G|$ then G does not have a hamiltonian square cycle, but then also does not satisfy $\delta(G) \geq \frac{2}{3}|G|$. Moreover if G has an almost independent set of size almost $\frac{1}{3}|G|$ and $\delta(G) \geq \frac{2}{3}|G|$, then we will see that G does have a hamiltonian square cycle.) Next they proved that if $|G|$ is sufficiently large, $\delta(G) \geq \frac{2}{3}|G|$, and G has an extremal configuration, then G has a hamiltonian square cycle. When there are no extremal configurations, the Regularity Lemma imposes a pseudo random structure on the graph that can be exploited, using this lack of extremal configurations and the Blow-up Lemma, to construct a hamiltonian square cycle. The use of the Regularity Lemma causes the constant n_0 to be extremely large.

Very recently Rödl, Ruciński and Szemerédi have made another important advance [16, 17]. They proved the following version of Dirac's Theorem for 3-uniform hypergraphs (3-graphs). An *open chain* $P := v_1v_2v_3 \dots v_{s-2}v_{s-1}v_s$ in a 3-graph H is a sequence of distinct

[1]The term reservoir is not mentioned in [6], and the modifiers *weak*, *strong* and *special* are our own invention. However, in light of more recent papers this terminology provides a consistent transition (see Definition 14).

vertices such that $v_i v_{i+1} v_{i+2} \in E(H)$ for all $i \in [s-2]$; P is a *closed chain* if in addition $v_{s-1}v_s v_1, v_s v_1 v_2 \in E(H)$.

Theorem 6 (Rödl, Ruciński and Szemerédi [17]). *There exists an integer n_0 such that for every 3-graph H on at least n_0 vertices, if every pair of vertices of H is contained in at least $\lfloor \frac{1}{2}|H| \rfloor$ edges of H then H contains a hamiltonian closed chain.*

The remarkable proof is very long, but has a similar structure to the proof of Theorem 5. However, a major difference is that the non-extremal case does not use any version of the Blow-up Lemma, and regularity (weak hypergraph regularity) is only used in a quite generic way to construct various *strong reservoirs*—weak reservoirs with no extreme sets. The Blow-up Lemma is replaced by a construction based on an ingenious *absorbing path* lemma, and a *connecting* lemma, that uses the strong reservoir.

Levitt, Sárközy and Szemerédi [9] applied similar techniques to the non-extremal case of Pósa's Conjecture without using the Regularity Lemma, and thus proved the result for much smaller graphs than those considered in Theorem 5.

Here we show that Pósa's Conjecture holds for graphs of order at least 2×10^8 without using the Regularity-Blow-up method. In addition, our proof of the extremal case holds for all n. We were influenced by the ideas of [9], but only rely on results from [6, 7, 8], and the idea from [12] of dividing the problem into an extremal case and a non-extremal case. We avoid the Blow-up Lemma and absorbing paths by using Theorem 4. Our approach is explained fully in the next section.

Notation

Most of our notation is consistent with Diestel's graph theory text [3]. In particular note that P^n is a path on n edges, $|G| = |V(G)|$, $\|G\| = |E(G)|$, and $d(v)$ is the degree of the vertex v. For $A, B \subseteq V(G)$, let $\|A, B\| = |E(A, B)|$, where $E(A, B)$ is the set of edges with one end in A and the other in B, in particular we shall write $\|a, B\|$ if $A = \{a\}$. We also use $\overline{\|A, B\|}$ to denote the number of edges in the complement of G that have one end in A and the other in B. For $a_1, a_2, \ldots, a_k \in V(G)$, let $N(a_1, a_2, \ldots, a_k) = N(a_1) \cap N(a_2) \cdots \cap N(a_k)$.

2 Main theorem and proof strategy

Here is our main result:

Theorem 7. *Let G be a graph on n vertices with $n \geq n_0 := 2 \times 10^8$. If $\delta(G) \geq \frac{2}{3}n$, then G has a hamiltonian square cycle.*

In this section we organize the structure of the proof. The first step is to define a usable extremal configuration. Our choice is simpler than the choice in [9], which was much simpler than the several extremal configurations used in [12]. A priori, this makes the extremal case easier and the non-extremal case harder.

Definition 8. Let G be a graph on n vertices. A set $S \subseteq V(G)$ is *α-extreme* if $|S| \geq (1-\alpha)\frac{n}{3}$ and $\|v, S\| < \alpha\frac{n}{3}$ for all $v \in S$.

The proof divides into two parts, depending on whether G is $\frac{1}{36}$-extreme, i.e., contains an α-extreme set with $\alpha := \frac{1}{36}$. The extreme case is handled in Section 4, where we prove the following theorem without assuming anything about the order of G. Its proof only requires elementary graph theory. Notice that $K_{3t+2} - E(K_{t+1})$ demonstrates that the degree condition is tight.

Theorem 9 (Extremal Case)**.** *Let G be a graph on n vertices with $\delta(G) \geq \frac{2}{3}n$. If G has a $\frac{1}{36}$-extreme set, then G has a hamiltonian square cycle.*

The non-extremal case is more complicated. In Section 3 we will prove:

Theorem 10 (Non-extremal Case)**.** *Let G be a graph on n vertices with $\delta(G) \geq \frac{2}{3}n$ and $n \geq n_0 := 2 \times 10^8$. If G does not contain a $\frac{1}{36}$-extreme set, then G has a hamiltonian square cycle.*

Note that if G has an α-extreme set $S \subseteq V(G)$ for some $\alpha < \frac{1}{36}$, then S is a $\frac{1}{36}$-extreme set. This explains why we only consider $\frac{1}{36}$-extreme sets in Theorems 9 and 10.

The proof of Theorem 10 has three parts. First we use the Reservoir Lemma (Lemma 22) to construct a special reservoir R with $|R| < \frac{1}{3}n$. Then we use the Path Cover Lemma (Lemma 23) to construct two disjoint square paths P_1, P_2 in $G - R$ such that $|P_1| + |P_2| > \frac{2}{3}n$ using techniques and results from [6, 7]. Finally, we use the properties of the special reservoir R, together with our version of the Connecting Lemma (Lemma 21), to connect the ends of P_1 to the ends of P_2 by disjoint square paths in R so as to form a square cycle of length greater than $\frac{2}{3}n$. Thus by Theorem 4 we obtain a hamiltonian square cycle.

2.1 Reservoirs and the Connecting Lemma

The bottleneck in this line of attack is in determining properties for special reservoirs that are strong enough to prove the Connecting Lemma, yet weak enough to ensure the existence of special reservoirs in moderately sized graphs. In the process of constructing a connecting square path we need to know that certain subsets of the reservoir are nonextreme. Since it is too expensive to ensure that all subsets are nonextreme, we anticipate a limited collection of *special* subsets that might appear in this construction, and construct a reservoir with no extreme special sets.

Definition 11. A set $S \subseteq V(G)$ is *special* if there exist (not necessarily distinct) vertices $u, v, w, x, y \in V(G)$ such that $S = (N(u, v, w) \cup N(u, v, x)) \cap N(y)$.

A set S of size at least $(1-\alpha)\frac{n}{3}$ that is not α-extreme has at least one vertex with "large" degree to S, but we will need more than one vertex of "large" degree, so we define a more general notion of extremity.

Definition 12. Let G be a graph with n vertices. A set $S \subseteq V(G)$ is (α, β)-*extreme* if $|S| \geq (1 - \alpha + \beta)\frac{n}{3}$ and there are fewer than $\lfloor \beta \frac{n}{3} \rfloor$ vertices $v \in S$ such that $\|v, S\| \geq \alpha \frac{n}{3}$.

So a set S of size at least $(1 - \alpha + \beta)\frac{n}{3}$ that is not (α, β)-extreme has at least $\lfloor \beta \frac{n}{3} \rfloor$ vertices with "large" degree to S. In the non-extremal case we know that G contains no α-extreme sets, but we must ensure for the Connecting Lemma that the reservoir has no (α', β')-extreme special sets. So we use the following simple observation when constructing the reservoir.

Lemma 13. *Let G be a graph on n vertices and let $\alpha, \beta > 0$. If G has no α-extreme sets and $S \subseteq V(G)$ with $|S| \geq (1 - \alpha + \beta)\frac{n}{3}$, then S is not (α, β)-extreme.*

Proof. Suppose S is (α, β)-extreme and let $S' = \{v \in S : \|v, S\| \geq \alpha \frac{n}{3}\}$. Since S is (α, β)-extreme, we have $|S'| < \lfloor \beta \frac{n}{3} \rfloor$. Thus $|S \setminus S'| \geq (1 - \alpha)\frac{n}{3}$ and $\|v, S \setminus S'\| < \alpha \frac{n}{3}$ for all $v \in S \setminus S'$, contradicting the fact that G has no α-extreme sets. □

Here are the technical definitions of (ϵ, ϱ)-weak, $(\alpha, \epsilon, \varrho)$-strong and $(\alpha, \beta, \epsilon, \varrho)$-special reservoir.

Definition 14 (Reservoir)**.** Let G be a graph on n vertices. Let $1 \geq \varrho \geq 0$ and $\epsilon > 0$. An (ϵ, ϱ)-*weak reservoir* is a set $R \subseteq V(G)$ such that $|R| = \lceil \varrho n \rceil$ and for all $u \in V(G)$,

$$\left(\frac{d(u)}{n} - \epsilon \right) |R| \leq \|u, R\| \leq \left(\frac{d(u)}{n} + \epsilon \right) |R|.$$

An $(\alpha, \epsilon, \varrho)$-*strong reservoir* is an (ϵ, ϱ)-weak reservoir R such that $G[R]$ has no α-extreme sets.

An $(\alpha, \beta, \epsilon, \varrho)$-*special reservoir* is an (ϵ, ϱ)-weak reservoir R such that for all special sets $S \subseteq V(G)$, $S \cap R$ is not (α, β)-extreme in $G[R]$.

A routine application of Chernoff's bound yields (ϵ, ϱ)-weak reservoirs R in moderately large graphs. The reason for this is that we have only polynomially many conditions to preserve. A similar observation allows us to construct $(\alpha, \beta, \epsilon, \varrho)$-special reservoirs. However this standard approach fails for $(\alpha, \epsilon, \varrho)$-strong reservoirs, because there are exponentially many conditions to check.

A connecting lemma should state that any two disjoint ordered edges in $V(G) \setminus R$ can be connected by a short square path whose interior vertices are in R. For example, Fan and Kierstead [6] proved:

Lemma 15. *If $\delta(G) > \frac{2}{3}|G|$ then there exists a square path connecting any two disjoint edges.*

In the context of Theorem 2, $(\epsilon/2, \varrho)$-weak reservoirs are sufficient since the degree bounds ensure that $\delta(G[R]) > \frac{2}{3}|R|$. In [9, 17] the authors prove connecting lemmas for strong reservoirs. We use a simpler argument and show that it works for special reservoirs.

2.2 Optimal paths

Let $e_1 := v_1v_2$ and $e_2 := v_{s-1}v_s$ be disjoint ordered edges. A square (e_1, e_2)-*path* is a square path of the form $v_1v_2 \ldots v_{s-1}v_s$.

Definition 16. An *optimal* square path (or cycle, or (e_1, e_2)-path) is a square path (or cycle, or (e_1, e_2)-path) P such that among all square paths (or cycles, or (e_1, e_2)-paths) (i) P is as long as possible, (ii) subject to (i), P has as many 3-chords as possible, and (iii) subject to (i) and (ii), P has as many 4-chords as possible.

All the work in [6, 7, 8] starts with lemmas about optimal square paths.

Lemma 17 (Fan-Kierstead [6], [7] Lemma 1)**.** *Suppose that P is a square path in a graph G and $v \in V(G-P)$. If P is an (e_1, e_2)-optimal square path then $\|v, Q\| \leq \frac{2}{3}|V(Q)| + 1$ for every segment Q of P. Moreover, if P is an optimal square path then $\|v, P\| \leq \frac{2}{3}|P| - \frac{1}{3}$ and if P is an optimal square cycle then $\|v, P\| \leq \frac{2}{3}|P| + \frac{1}{3}$.*

In the extremal case we will take advantage of the following fact.

Corollary 18. *Pósa's Conjecture is true, if it holds for all G with $|G|$ divisible by* 3.

Proof. Suppose $|G| = 3k + r$, where $1 \leq r \leq 2$. Let G' be G with r vertices deleted. Then

$$\delta(G') \geq \lceil \frac{2}{3}(3k + r) \rceil - r = 2k = \frac{2}{3}|G'|.$$

Thus by hypothesis, G' has a hamiltonian square cycle C'. So an optimal square cycle C in G has length at least $3k$. Suppose C is not hamiltonian in G. Then there exists $x \in V(G - C)$. By Lemma 17, we have the following contradiction:

$$2k + r \leq \delta(G) \leq \|v, C\| + |G| - |C| - 1 \leq |G| - \frac{1}{3}|C| - \frac{2}{3} \leq 2k + r - \frac{2}{3}.$$

□

We will also need:

Lemma 19 (Fan-Kierstead [7], Lemma 9)**.** *Let P be an optimal square path of G. Let xy be an edge of $G - P$ such that there are square paths, of at least q vertices, starting at xy and yx in $G - P$. If $|P| \geq 2q + 2$, then $\|xy, P\| \leq \frac{4}{3}|P| - \frac{2}{3}q + 2$.*

2.3 Probability

If X is a random variable with hypergeometric distribution (and our experiment consists of drawing n items from a collection of N total items, m of which are good and $N - m$ of which are bad) the expected value of X is given by

$$\mathbb{E}X = \sum_{k=0}^{n} k \cdot Pr(X = k) = \sum_{k=0}^{n} k \cdot \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}} = \frac{nm}{N}.$$

Theorem 20 (Chernoff's bound [2, 10])**.** *Let X be a random variable with binomial or hypergeometric distribution. Then the following hold:*

1. $Pr(X \geq \mathbb{E}X + t) \leq \exp\left(-\frac{t^2}{2(\mathbb{E}X + t/3)}\right)$, $\quad t \geq 0$;
2. $Pr(X \leq \mathbb{E}X - t) \leq \exp\left(-\frac{t^2}{2\mathbb{E}X}\right)$, $\quad t \geq 0$;
3. If $0 < \gamma \leq 3/2$, then $Pr(|X - \mathbb{E}X| \geq \gamma \mathbb{E}X) \leq 2\exp\left(-\frac{\gamma^2}{3}\mathbb{E}X\right)$.

3 Non-extremal case

In this section we prove Theorem 10. We have compromised optimality somewhat in our constructions and calculations in favor of clarity of exposition. For instance, we know how to reduce n_0 by a factor of 2. That being said, we can make the reservoir lemma slightly simpler and we can choose "nicer" constants throughout the non-extremal case at the cost of a factor of 3 in n_0.

We first show that if H is a graph with no (α, β)-extreme special sets whose minimum degree is almost $\frac{2}{3}|H|$, then any two disjoint edges in H can be connected by a short square path. Let $xy \in E(H)$; we say that $P\{xy\}Q$ is a square path if one of $PxyQ$ or $PyxQ$ is a square path.

Lemma 21 (Connecting Lemma)**.** *Let $0 < \beta < \alpha \leq \frac{1}{36}$, $0 < \epsilon \leq \frac{\alpha-\beta}{15.1}$, $l := 10$ and suppose $n \geq \max\{\frac{660}{\epsilon}, \frac{69}{\beta}\}$. Let $H = (V, E)$ be a graph on n vertices with no (α, β)-extreme special sets such that $\delta(H) \geq (\frac{2}{3} - \epsilon)n$. Let $L \subseteq V$ such that $|L| \leq l$. If ab, cd are any two disjoint ordered edges in $H - L$, then there is a square (ab, cd)-path P of order at most 14 for which $V(P) \subseteq V \setminus L$.*

Proof. Let ab, cd be disjoint ordered edges in $H - L$ and set $A := \{a, b, c, d\}$. Here is our plan. First (a) we find disjoint edges $a'b', c'd'$ in $H - L - A$ such that $\|ab, a'b'\| = 4 = \|cd, c'd'\|$. Then, setting $A' := \{a', b', c', d'\}$, (b) we construct a square path $\{a'b'\}Q\{c'd'\}$ with $Q \subseteq H' := H \setminus (L \cup A \cup A')$ connecting the unordered edges $a'b', c'd'$. This will yield a square path $ab\{a'b'\}Q\{c'd'\}cd$, where the order of $\{a'b'\}$ and $\{c'd'\}$ is determined by Q.

Let $M \subseteq V$ with $|M| \leq l + 12$. We will often use the following statement:

$$\text{If } S \text{ is a special set with } |S| \geq (1 - \alpha + \beta)\frac{n}{3} \text{ then } \|S \setminus M\| > 0. \tag{1}$$

To see this, note that since S is not (α, β)-extreme and $n \geq \frac{69}{\beta}$, S has at least $\left\lfloor \beta\frac{n}{3} \right\rfloor > l + 12$ vertices with degree at least $\alpha\frac{n}{3} > l + 12$.

Consider the special set $N(a, b) = (N(a, a, a) \cup N(a, a, a)) \cap N(b)$. Since $\delta(H) \geq (\frac{2}{3} - \epsilon)n$, we have

$$|N(a, b)| \geq (1 - 6\epsilon)\frac{n}{3} \geq (1 - \alpha + \beta)\frac{n}{3}.$$

By (1), there exists $a'b' \in E(N(a,b) \setminus (L \cup A))$. Likewise there is an edge $c'd' \in E(N(c,d) \setminus (\{a',b'\} \cup L \cup A))$, completing the first goal (a).

Next we show (b). Let $V' := V(H')$. Then $|V'| \geq n - l - 8$. We must construct $Q \subseteq H'$. For $i \in [4]$, let $S_i := S_i(A') = \{v \in V : \|v, A'\| = i\}$. Then

$$\frac{8}{3}n - 4\epsilon n = 4(\frac{2}{3} - \epsilon)n \leq \|A', V\| = \sum_{i \in [4]} i|S_i| \leq 4|S_4| + 3|S_3| + 2(n - |S_4| - |S_3|), \tag{2}$$

which gives

$$2|S_4| + |S_3| \geq \frac{2}{3}n - 4\epsilon n. \tag{3}$$

Case 1: $|S_4| > l + 12$. Looking ahead to an application in Case 2.a, we will construct $Q \subseteq H'' := H' - A''$, for any fixed 4-set A''. Set $V'' := V(H'')$. By the case assumption, there exists $x \in S_4 \cap V''$. If there exists $u \in N(x) \cap (S_4 \cup S_3) \cap V''$ then set $Q := \{xu\}$. Otherwise, $|S_4| + |S_3| \leq \frac{1}{3}n + \epsilon n + l + 12$, since $d(x) \geq \frac{2}{3}n - \epsilon n$. Thus by (3), and using $\alpha - \beta \geq 15.1\epsilon$ and $n \geq \frac{660}{\epsilon}$, we have

$$|S_4| \geq \frac{1}{3}n - 5\epsilon n - l - 12 \geq (1 - \alpha + \beta)\frac{n}{3}.$$

Moreover, $S_4 = N(a',b',c',d') = (N(a',b',c') \cup N(a',b',c')) \cap N(d')$ is special. Thus by (1), there exists an edge $uv \in S_4 \cap V''$, and we set $Q := uv$.
Case 2: $|S_4| \leq l + 12$. Let

$$T_1 := \{v \in S_3 \cup S_4 : \|v, \{a',b'\}\| = 2\} = (N(a',b',c') \cup N(a',b',d')) \cap N(a') \text{ and} \tag{4}$$
$$T_2 := \{v \in S_3 \cup S_4 : \|v, \{c',d'\}\| = 2\} = (N(c',d',a') \cup N(c',d',b')) \cap N(c'). \tag{5}$$

Then T_1 and T_2 are both special sets. Note that S_3 is partitioned as $(T_1 \setminus S_4) \cup (T_2 \setminus S_4)$ and $T_1 \cap T_2 = S_4$. By (3) and the fact that $|T_1| + |T_2| = |S_3| + 2|S_4|$, we have

$$|T_1| + |T_2| \geq \frac{2}{3}n - 4\epsilon n. \tag{6}$$

Without loss of generality, $|T_1| \leq |T_2|$, and so $T_2 \neq \emptyset$. Finally, note that by (3) and the case assumption we have,

$$|T_1 \cup T_2| = |S_3 \cup S_4| \geq \frac{2}{3}n - 4\epsilon n - l - 12. \tag{7}$$

Case 2.a: $|T_1| > l + 8$. If there exists $xy \in E(T_1, T_2) \cap E(H')$, then set $Q := xy$. Otherwise, let $x \in T_1 \cap V'$. Then using, in order, $d(x) \geq (\frac{2}{3} - \epsilon)n$, (6), $\alpha - \beta \geq 15.1\epsilon$ and $n \geq \frac{660}{\epsilon}$ we have

$$\frac{n}{3} + \epsilon n + l + 8 \geq |T_2| \geq |T_1| \geq \frac{n}{3} - 5\epsilon n - l - 8 \geq (1 - \alpha + \beta)\frac{n}{3}. \tag{8}$$

By (1) and (8), there exist edges $a''b'' \in E(T_1)$ and $c''d'' \in E(T_2)$ such that $A'' := \{a'',b'',c'',d''\}$ is disjoint from $L \cup A \cup A'$. Note that $A'' \cap S_4 = \emptyset$, since $E(T_1, T_2) \cap E(H') = \emptyset$ as mentioned above.

Set $U := V \setminus (T_1 \cup T_2)$. By (7),

$$|U| = n - |T_1 \cup T_2| \leq \frac{n}{3} + 4\epsilon n + l + 12. \tag{9}$$

By (8), for any $x \in A''$,

$$\|x, U\| \geq \frac{2}{3}n - \epsilon n - |T_2| \geq \frac{n}{3} - 2\epsilon n - l - 8. \tag{10}$$

By (9), (10), and $n \geq \frac{660}{\epsilon}$, we have $\overline{\|x, U\|} \leq 6\epsilon n + 3l + 32 < \frac{1}{5}|U \cap V''|$. Thus there exist more than $l + 12$ vertices in $S_4(A'')$. Thus by Case 1, there exists a square path $Q := \{a''b''\}Q'\{c''d''\}$ with $|Q'| \leq 2$.
Case 2.b: $|T_1| \leq l + 8$. Then $|T_2| \geq \frac{2}{3}n - 4\varepsilon n - l - 8$ by (6). Let $x \in N(a', b') \cap V'$, and note that $S := T_2 \cap N(x) = (N(a', c', d') \cup N(b', c', d')) \cap N(x)$ is a special set. Moreover by $\alpha - \beta \geq 15.1\epsilon$ and $n \geq \frac{660}{\epsilon}$ we have

$$|S| \geq |T_2| + |N(x)| - n \geq \frac{n}{3} - 5\epsilon n - l - 8 \geq (1 - \alpha + \beta)\frac{n}{3}.$$

Thus by (1), there exists an edge $yz \in E(S \cap V')$. Let $Q := xyz$.

$\square$

Now we prove the reservoir lemma.

Lemma 22 (Reservoir Lemma)**.** *Let $\alpha \geq \frac{1}{36}$, $c \geq \frac{1}{14}$, $\alpha' := (1 - 3c)\alpha$, $\beta' := c\alpha$, $\epsilon \geq \frac{\alpha' - \beta'}{15.1}$, $\varrho \geq 1 - \frac{2/3+\epsilon}{5/6-2\epsilon}$ and $n \geq n_0 := 2 \times 10^8$. If H is a graph on n vertices such that $\delta(H) \geq \frac{2}{3}n$ and H contains no α-extreme sets, then H contains an $(\alpha', \beta', \epsilon, \varrho)$-special reservoir.*

Proof. Let $\gamma := \frac{2\beta'}{1-\alpha'-\beta'}$. We will show that there exists a set $R \subseteq V(H)$ such that $|R| = \lceil \varrho n \rceil$ which satisfies the following three properties.

(i) For all $u \in V(H)$, $\left(\frac{d(u)}{n} - \epsilon\right)|R| \leq \|u, R\| \leq \left(\frac{d(u)}{n} + \epsilon\right)|R|$.

(ii) For all special sets $S \subseteq V(H)$, if $|S| \geq (1 - \alpha' + \beta')\varrho\frac{n}{3}$, then $|S \cap R| \leq 1.05\varrho|S|$ and for all special sets $S \subseteq V(H)$, if $|S \cap R| \geq (1 - \alpha' + \beta')\varrho\frac{n}{3}$, then $|S \cap R| \leq (1 + \gamma)\varrho|S|$.

(iii) For all special sets $S \subseteq V(H)$, if $|S| \geq (1 - \alpha' - \beta')\frac{n}{3}$, then there exists a set $T' \subseteq S \cap R$ such that $|T'| \geq \beta'\varrho\frac{n}{3}$ and $\|z, S \cap R\| \geq \alpha'\varrho\frac{n}{3}$ for all $z \in T'$.

Then we will show that these three properties imply that R is an $(\alpha', \beta', \epsilon, \varrho)$-special reservoir.

Let $R \subseteq V(H)$ be a set of size $\lceil \varrho n \rceil =: r$ chosen at random from all $\binom{n}{r}$ possibilities. There are five calculations that follow. In each of these calculations we will need n to be large, specifically $n \geq 2 \times 10^8$ is large enough.

Let $u \in V(H)$. The expected value of $\|u, R\|$ is $\frac{rd(u)}{n} \geq \varrho d(u)$. So by Theorem 20.3, we have

$$Pr\left(\left|\|u, R\| - \frac{rd(u)}{n}\right| \geq \frac{\epsilon n}{d(u)}\frac{rd(u)}{n}\right) \leq 2\exp\left(-\frac{\left(\frac{\epsilon n}{d(u)}\right)^2}{3}\frac{rd(u)}{n}\right) \leq 2\exp\left(\frac{-\epsilon^2\varrho n^2}{3d(u)}\right) < \frac{1}{3n}.$$

There are n vertices in $V(H)$. So by applying Boole's inequality, the probability that there exists a vertex which does not satisfy property (i) is less than 1/3.

Let $S \subseteq V(H)$ be a special set such that $|S| \geq (1-\alpha'+\beta')\varrho\frac{n}{3}$. The expected value of $|S \cap R|$ is $\frac{r|S|}{n} \geq \varrho|S| \geq (1-\alpha'+\beta')\varrho^2\frac{n}{3}$. So by Theorem 20.1, we have

$$\log Pr(|S\cap R| \geq 1.05\frac{r|S|}{n}) \leq -\frac{(.05\varrho|S|)^2}{2(\varrho|S|+.05\varrho|S|/3)} \leq -\frac{.0025\varrho^2(1-\alpha'+\beta')}{2(1+.05/3)}\frac{n}{3} < \log\frac{1}{9n^5}.$$

So with high probability,

$$|S\cap R| \leq 1.05\varrho|S| \text{ for all } S\subseteq V(H) \text{ such that } |S| \geq (1-\alpha'+\beta')\varrho\frac{n}{3}. \tag{11}$$

Now let $S \subseteq V(H)$ be a special set such that $|S\cap R| \geq (1-\alpha'+\beta')\varrho\frac{n}{3}$. Since $|S| \geq |S\cap R|$ we have $|S| \geq (1-\alpha'+\beta')\varrho\frac{n}{3}$ and thus by (11), $|S| \geq \frac{|S\cap R|}{1.05\varrho} \geq \frac{(1-\alpha'+\beta')}{1.05}\frac{n}{3}$. The expected value of $|S\cap R|$ is $\frac{r|S|}{n} \geq \varrho|S| \geq \varrho\frac{(1-\alpha'+\beta')}{1.05}\frac{n}{3}$. Using Theorem 20.1 again, we have

$$\log Pr(|S\cap R| \geq (1+\gamma)\frac{r|S|}{n}) \leq -\frac{(\gamma\varrho|S|)^2}{2(\varrho|S|+\gamma\varrho|S|/3)} \leq -\frac{\gamma^2\varrho(1-\alpha'+\beta')}{1.05(2+2\gamma/3)}\frac{n}{3} < \log\frac{1}{3n^5}.$$

There are at most n^5 special sets $S\subseteq V(H)$. So by applying Boole's inequality, the probability that there exists a set S which does not satisfy property (ii) is less than 4/9.

Let $S \subseteq V(H)$ be a special set such that $|S| \geq (1-\alpha'-\beta')\frac{n}{3} = (1-\alpha+2c\alpha)\frac{n}{3}$. Since H has no α-extreme sets, we see by Lemma 13 that S is not $(\alpha, 2c\alpha)$-extreme. So there exists a set $S' \subseteq S$ having the property that $|S'| = \lfloor 2c\alpha\frac{n}{3}\rfloor$ and for all $v\in S'$, $\|v,S\| \geq \alpha\frac{n}{3}$. Let $T' := S'\cap R$. We first show that with high probability, $|T'| \geq \frac{3\varrho}{4}|S'| \geq \frac{\varrho}{2}(|S'|+1) \geq \beta'\varrho\frac{n}{3}$. The expected value of $|T'|$ is $\varrho|S'| \geq \varrho(2c\alpha\frac{n}{3}-1)$. So by Theorem 20.2, we have

$$\log Pr(|T'| \leq \varrho|S'| - \frac{\varrho}{4}|S'|) \leq -\frac{(\frac{\varrho}{4}|S'|)^2}{2(\varrho|S'|)} = -\frac{\varrho|S'|}{32} \leq -\frac{\varrho(2c\alpha\frac{n}{3}-1)}{32} < \log\frac{1}{9n^5}.$$

Next we show that, with high probability, every vertex in S' has at least $(1-3c)\varrho\|v,S\| \geq \alpha'\varrho\frac{n}{3}$ neighbors in $S\cap R$. Let $v\in S'$. The expected value of $\|v,T\|$ is $\varrho\|v,S\| \geq \varrho\alpha\frac{n}{3}$. So by Theorem 20.2, we have

$$\log Pr(\|v,S\cap R\| \leq (1-3c)\varrho\|v,S\|) \leq -\frac{(3c\varrho\|v,S\|)^2}{2\varrho\|v,S\|} = -\frac{9c^2\varrho\|v,S\|}{2} \leq -\frac{3c^2\varrho\alpha n}{2} < \log\frac{1}{9n^6}.$$

There are at most n^5 special sets $S\subseteq V(H)$ and at most n^6 sets defined when we examine the neighborhood of vertices in each special set. So by applying Boole's inequality, the probability that there exists a set S which does not satisfy property (iii) is less than 2/9.

The probability that R doesn't satisfy one of the conditions is less than 1, thus there exists a set $R\subseteq V(H)$ satisfying properties (i)-(iii).

We now show that R is an $(\alpha', \beta', \epsilon, \varrho)$-special reservoir. Since R satisfies property (i), R is a (ϵ, ϱ)-weak reservoir. Let $S \subseteq V(H)$ be a special set such that $|S \cap R| \geq (1-\alpha'+\beta')\varrho\frac{n}{3}$. By property (ii), we have $\varrho|S|(1+\gamma) \geq |S \cap R| \geq (1-\alpha'+\beta')\varrho\frac{n}{3}$, and thus

$$|S| \geq \frac{(1-\alpha'+\beta')}{1+\gamma}\frac{n}{3} = (1-\alpha'-\beta')\frac{n}{3}.$$

Then since $|S| \geq (1-\alpha'-\beta')\frac{n}{3}$ there is, by property (iii), a set of vertices $T' \subseteq S \cap R$ with $|T'| \geq \beta'\varrho\frac{n}{3}$ such that for all $v \in T'$, $\|v, S \cap R\| \geq \alpha'\varrho\frac{n}{3}$. Thus $S \cap R$ is not (α', β')-extreme in $G[R]$. Therefore R is an $(\alpha', \beta', \epsilon, \varrho)$-special reservoir. □

We now prove a lemma which allows us to cover most of the complement of the reservoir with at most two long square paths.

Lemma 23 (Path Cover Lemma). *Suppose $\epsilon \leq \frac{1}{500}$ and $n \geq 6000$. Let H be a graph on n vertices with $\delta(H) \geq \left(\frac{2}{3}-\epsilon\right)n$. Then*

(a) H has a square path P with $|P| \geq (\frac{1}{2}-3\epsilon)n$.

(b) H has two vertex disjoint square paths P_1 and P_2 so that $|P_1|+|P_2| > (\frac{5}{6}-2\epsilon)n$.

Proof. (a) Let $P := u_1u_2...u_p$ be an optimal square path in H and suppose that $p < (\frac{1}{2}-3\epsilon)n$. We first observe that since $\delta(H) \geq (\frac{2}{3}-\epsilon)n$ we have $N(u_1,u_2) \geq (\frac{1}{3}-2\epsilon)n$ and thus $p > (\frac{1}{3}-2\epsilon)n$. Let $H' := H - P$ and set $h := |H'|$. If $\|v, P\| \leq (\frac{2}{3}-4\epsilon)p$ for all $v \in V(H')$ then we have $\delta(H') \geq (\frac{2}{3}-\epsilon)n - (\frac{2}{3}-4\epsilon)p \geq \frac{2}{3}h$. Thus by Theorem 3, H' has a hamiltonian square path of length more than than $\frac{1}{2}n$, contradicting the optimality of P. Thus there is a vertex $x \in V(H')$ such that $\|x, P\| > (\frac{2}{3}-4\epsilon)p > \frac{1}{2}p+1$. It follows that x is adjacent to two consecutive vertices of P. Choose $i \in [p]$ as small as possible such that $u_i, u_{i+1} \in N(x)$. Let $Q := u_1u_2...u_{i-1}$ and set $q := i-1$. Then $\|x, Q\| \leq \frac{1}{2}q$. We claim that $q < (\frac{1}{6}-2\epsilon)n$. Otherwise,

$$\begin{aligned}\|x, P-Q\| > (\frac{2}{3}-4\epsilon)p - \frac{1}{2}q &= \frac{2}{3}(p-q) + \frac{1}{6}q - 4\epsilon p \\ &> \frac{2}{3}|P-Q| + \frac{1}{6}(\frac{1}{6}-2\epsilon)n - 4\epsilon(\frac{1}{2}-3\epsilon)n \\ &> \frac{2}{3}|P-Q| + \frac{1}{36}n - \frac{7}{3}\epsilon n \\ &> \frac{2}{3}|P-Q| + 1,\end{aligned}$$

contradicting Lemma 17. On the other hand, since $|N(x,u_i)| \geq (\frac{1}{3}-2\epsilon)n = \frac{2}{3}(\frac{1}{2}-3\epsilon)n > \frac{2}{3}p$, Lemma 17 implies x and u_i have a common neighbor y in H'. Also, by Lemma 17 we have

$$\delta(H') \geq (\frac{2}{3}-\epsilon)n - (\frac{2}{3}p - \frac{1}{3}) > \frac{2}{3}h - \epsilon n,$$

and thus for any edge uv in H', $|N_{H'}(u,v)| \geq \frac{1}{3}h - 2\epsilon n > (\frac{1}{6}-2\epsilon)n$. Hence, we can find a square path P' of length at least $(\frac{1}{6}-2\epsilon)n$ starting at xy. Since $|P'| > q$, the square path $P'yxu_iu_{i+1}...u_p$ is longer than P, a contradiction. This completes the proof of part (a).

(b) Let P_1 be an optimal square path in H and let $p := |P_1|$. Note that $p \geq (\frac{1}{2} - 3\epsilon)n$ by Lemma 23(a). If $p > (\frac{5}{6} - 2\epsilon)n$, then set $P_2 = \emptyset$ and we are done. So we may assume that $p \leq (\frac{5}{6} - 2\epsilon)n$. Set $H' := H - P_1$ and $h := |H'| > n/6$. If $\|v, P_1\| \leq (\frac{2}{3} - 3\epsilon)p$ for all $v \in V(H')$ then $\delta(H') \geq (\frac{2}{3} - \epsilon)n - (\frac{2}{3} - 3\epsilon)p \geq \frac{2}{3}h$. Thus H' has a hamiltonian square path P_2 by Theorem 3, and we are done. Otherwise, let $x \in V(H')$ such that $\|x, P_1\| > (\frac{2}{3} - 3\epsilon)p$. Note that by Lemma 17, we have $\delta(H') \geq (\frac{2}{3} - \epsilon)n - (\frac{2}{3}p - \frac{1}{3}) > \frac{2}{3}h - \epsilon n$, and thus there is a square path of length at least $\frac{1}{3}h - 2\epsilon n$ starting at any ordered edge in H'. Set $H'' := G[N_{H'}(x)]$ and $h' := |H''|$. Note that by Lemma 19, we have that for all $y \in V(H'')$,

$$\|y, P_1\| < \frac{4}{3}p - \frac{2}{3}(\frac{1}{3}h - 2\epsilon n) + 2 - (\frac{2}{3} - 3\epsilon)p = \frac{2}{3}p - \frac{2}{9}h + \frac{4}{3}\epsilon n + 3\epsilon p + 2,$$

so

$$\|y, H'\| > (\frac{2}{3} - \epsilon)n - (\frac{2}{3}p - \frac{2}{9}h + \frac{4}{3}\epsilon n + 3\epsilon p + 2) = \frac{8}{9}h - \frac{7}{3}\epsilon n - 3\epsilon p - 2.$$

So every vertex in H'' has at most $\frac{1}{9}h + \frac{7}{3}\epsilon n + 3\epsilon p + 1$ nonneighbors in H'. Therefore

$$\delta(H'') \geq \frac{\frac{2}{3}h - \epsilon n - (\frac{1}{9}h + \frac{7}{3}\epsilon n + 3\epsilon p + 1)}{\frac{2}{3}h - \epsilon n} h' > \frac{2}{3}h',$$

since $\epsilon \leq \frac{1}{500}$, $n \geq 6000$, and $h > n/6$. Therefore H'' has a hamiltonian square path P_2. Thus

$$|P_1| + |P_2| > p + \frac{2}{3}h - \epsilon n = n - \frac{1}{3}h - \epsilon n \geq n - \frac{1}{3}(\frac{1}{2} + 3\epsilon)n - \epsilon n = (\frac{5}{6} - 2\epsilon)n.$$

□

Now we are ready to finish the nonextreme case.

Proof of Theorem 10. Let $\alpha := \frac{1}{36}$ and let G be a graph on n vertices. Suppose G has no α-extreme sets, $n \geq n_0 := 2 \times 10^8$, and $\delta(G) \geq \frac{2}{3}n$. Let $c := \frac{1}{14}$, $\epsilon := \frac{50}{1057}\alpha$, and $\varrho := 1 - \frac{2/3+\epsilon}{5/6-2\epsilon}$. Apply Lemma 22 to obtain an $(\frac{11}{14}\alpha, \frac{1}{14}\alpha, \epsilon, \varrho)$-special reservoir R. Let $H := G - R$ and let $h := |H|$. Since R is a special reservoir we have $\delta(H) \geq (\frac{2}{3} - \epsilon)h$. Now we apply Lemma 23 to H, to get disjoint square paths P_1 and P_2 so that

$$|P_1| + |P_2| > (\frac{5}{6} - 2\epsilon)h = (\frac{5}{6} - 2\epsilon)(n - \lceil \varrho n \rceil) \geq (\frac{2}{3} + \epsilon)n - 1 > \frac{2}{3}n.$$

Since R is a special reservoir, every special set $S \subseteq V(G)$ has the property that $S \cap R$ is not $(\frac{11}{14}\alpha, \frac{1}{14}\alpha)$-extreme in $G[R]$. So we apply Lemma 21 at most twice to connect the paths P_1 and P_2 through R. On the second application, we set $L := V(P_1) \cap R$ to make sure that we avoid the vertices used in the first application. This gives us a square cycle C with $V(P_1) \cup V(P_2) \subseteq V(C)$ and thus $|C| > \frac{2}{3}n$. Therefore G has a hamiltonian square cycle by Theorem 4. □

4 Extremal Case

In this section we prove Theorem 9. First we need two propositions. Note that the length of an (ordinary) path P is the size $\|P\|$ of its edge set.

Proposition 24. *Every connected graph H with $|H| \geq 3$ has a path or cycle of length* $\min(2\delta(H), |H|)$.

Proof. Let P be a maximum length path in H. If we are not done, then $\|P\| < 2\delta(H)$. So, as in the proof of Dirac's Theorem [4], G has a cycle C that spans $V(P)$. If C is hamiltonian then we are done; otherwise, using connectivity, we can extend C to a path longer than P, a contradiction. $\square$

Proposition 25. *If H is a graph with circumference $l > |H| - \delta(H)$, then $l \geq \min(2\delta(H), |H|)$, and moreover, if $|H|$ is also even, then H has an even cycle of length at least* $\min(2\delta(H), |H|)$.

Proof. Let $C \subseteq H$ be a cycle of length l, and fix an orientation of C. If $|C| = |H|$ then we are done, even if $|H|$ is even. Otherwise, let $P := v_1 \dots v_p$ be a maximum path in $H - C$. Then all neighbors of v_p are on $P \cup C$. By hypothesis $\delta(H) > |H| - l \geq p$, and so v_1 has a neighbor $x \in C$ and v_p has a neighbor on $C - x$. Let $y, z \neq x$ be neighbors of v_p on C with y as close as possible to x in the forward direction and z as close as possible in the backward direction (possibly $y = z$). Then $\|zCx\|, \|xCy\| \geq p + 1$, as otherwise we could replace the interior vertices of one of these segments with P to obtain a longer cycle, which would yield a contradiction. Moreover, since C has maximum length, any two neighbors of v_p are separated by at least one vertex on C. Since v_p has at least $d(v_p) - p$ neighbors on $C - x$,

$$|C| = \|xCy\| + \|yCz\| + \|zCx\| \geq (p+1) + 2(d(v_p) - p - 1) + (p+1) \geq 2\delta(H).$$

Now suppose $|H|$ is even. If $|C|$ is even we are done, so suppose $|C|$ is odd. Consider the path P and vertices x, y, z defined above. If $\|xCy\|$ and $\|zCx\|$ have different parity, then replace xCy with xPy or replace zCx with zPx to get an even cycle of length at least $2\delta(H)$. So assume $\|xCy\|$ and $\|zCx\|$ have the same parity, and thus $\|yCz\|$ is odd. Now v_p has $k \geq d(v_p) - p$ neighbors on yCz. Let $y = a_1, a_2, \dots, a_k = z$ be the neighbors of v_p on yCz in their natural order. Since $\|yCz\|$ is odd, some segment a_iCa_{i+1} must have odd length. By replacing a_iCa_{i+1} with $a_iv_pa_{i+1}$, we get a cycle C' with even length such that $|C'| \geq (p+1) + (p+1) + 2(d(v_p) - p - 1) \geq 2\delta(H)$ as before. $\square$

Proof of Theorem 9. Let $G = (V, E)$ be a graph on n vertices with $\delta(G) \geq \frac{2}{3}n$. By Corollary 18 we may assume $n = 3k$, which gives $\delta(G) \geq 2k$. Set $\alpha := \frac{1}{36}$, and suppose G has an α-extreme subset. Let $S \subseteq V$ be an α-extreme set of minimal order, so $|S| = \lceil (1-\alpha)k \rceil$. Set $T := V \setminus S$. If $k < 1/\alpha$, then $|S| = k$, $|T| = 2k$, $G[S, T]$ is complete and $\delta(G[T]) \geq k$. So by Dirac's theorem T has a hamiltonian cycle $C := y_1 \dots y_{2k} y_1$. Since $G[S, T]$ is complete we can insert the vertices $x_1, x_2, \dots, x_k$ of S into C so that $y_1y_2x_1y_3y_4x_2 \dots y_{2k-1}y_{2k}x_ky_1y_2$ is a hamiltonian square cycle. So for the rest of the proof assume $k \geq 1/\alpha$. Choose $T_0 \subseteq T$

such that $|V \setminus (S \cup T_0)|$ is even, $2\lfloor\sqrt{\alpha}k\rfloor - 1 \le |T_0| \le 2\lfloor\sqrt{\alpha}k\rfloor$, and subject to this, $\|T_0, S\|$ is as small as possible. Set $T_1 := T \setminus T_0$, and note that $|T_1|$ is even. We have,

$$\forall x \in S,\ \overline{\|x, T\|} \le k - (|S| - \|x, S\|) \le 2\lfloor\alpha k\rfloor. \tag{12}$$

Every vertex in T_1 has at most as many nonneighbors in S as every vertex in T_0. Thus, using $\alpha = \frac{1}{36}$, and expressing k as $k = 36q + r$ with $q, r \in \mathbb{Z}$ and $0 \le r \le 35$, we have

$$\forall y \in T_1,\ \overline{\|y, S\|} \le \left\lfloor \frac{2\lfloor\alpha k\rfloor |S|}{|T_0 \cup \{y\}|} \right\rfloor \le \left\lfloor \frac{2\lfloor\alpha k\rfloor (k - \lfloor\alpha k\rfloor)}{2\lfloor\sqrt{\alpha}k\rfloor} \right\rfloor \le \left\lfloor \frac{(35q + r)}{6} \right\rfloor \le \lfloor\sqrt{\alpha}k\rfloor. \tag{13}$$

Set $m := k - |T_0| + \lfloor\alpha k\rfloor$ and note that since $k \ge 36$,

$$m \ge \frac{2}{3}k + \lfloor\alpha k\rfloor \ge \frac{2}{3}k + 1. \tag{14}$$

Thus we have

$$\delta(G[T_1]) \ge 2k - |S \cup T_0| = k - |T_0| + \lfloor\alpha k\rfloor = m \ge \frac{2}{3}k + 1. \tag{15}$$

Case 1: There exists an even cycle $C \subseteq G[T_1]$ of length $2l \ge 2m$; say $C := y_1 \dots y_{2l} y_1$. Looking ahead to an application in Case 2, we prove something slightly more general than what is needed for Case 1. For some $t \le |T_1|/2$, let $T_1' \subseteq T_1$ such that $|T_1'| = 2t$. Enumerate the vertices of T_1' as $z_1, \dots, z_{2t}$. Let $P := \{p_1, \dots, p_t\}$ be a set of *ports*, where $p_i := \{z_{2i-1}, z_{2i}, z_{2i+1}, z_{2i+2}\}$ and addition of indices is modulo t. We say that a vertex $x \in S$ can be inserted into port p_i if $p_i \subseteq N(x)$.

Claim 26. *For $S' \subset S$ with $|S'| \ge |S| - 4$, let Γ be the S', P-bigraph with $xp \in E(\Gamma)$ if and only if x can be inserted into p. Then Γ has a matching $M := \{x_i p_i : i \in [t]\}$ that saturates P.*

Proof. Using Hall's Theorem, since $|S'| \ge |T_1|/2 \ge |P|$, it suffices to show that

$$\|x, P\|_\Gamma + \|S', p\|_\Gamma \ge |P| \text{ for all } x \in S' \text{ and } p \in P. \tag{16}$$

If $x \in S'$, then $\overline{\|x, T\|}_G \le 2\lfloor\alpha k\rfloor$ by (12). Since each $y \in T_1'$ is in two ports, each nonedge xy contributes to two nonedges in Γ. So $\overline{\|x, P\|}_\Gamma \le 4\lfloor\alpha k\rfloor$. Thus

$$\|x, P\|_\Gamma \ge |P| - \overline{\|x, P\|}_\Gamma \ge |P| - 4\alpha k. \tag{17}$$

If $p \in P$, then $\overline{\|S', y\|}_G \le \lfloor\sqrt{\alpha}k\rfloor$ for each $y \in p$ by (13). Thus $\overline{\|S', p\|}_\Gamma \le 4\lfloor\sqrt{\alpha}k\rfloor$. So

$$\|S', p\|_\Gamma \ge |S'| - \overline{\|S', p\|}_\Gamma \ge (1 - \alpha - \frac{4}{k} - 4\sqrt{\alpha})k. \tag{18}$$

Since $4\sqrt{\alpha} + 5\alpha + \frac{4}{k} \le \frac{33}{36} < 1$, summing (17) and (18) yields (16). $\square$

Let $S' := S$ and $P := \{p_1, \ldots, p_l\}$, where $p_i := \{y_{2i-1}, y_{2i}, y_{2i+1}, y_{2i+2}\}$ and addition of indices is modulo $2l$. By Claim 26, there exist $x_1, \ldots, x_l$ such that $y_1y_2x_1y_3y_4x_2 \ldots y_{2l-1}y_{2l}x_ly_1y_2$ is a square cycle of length $3l$. By (15), $3l \geq 3m > 2k$, and so Theorem 4 implies that G has a hamiltonian square cycle.

Case 2: Not Case 1. Since $|T_1|$ is even, using Proposition 25 and (15),

$$|D| \leq |T_1| - \delta(G[T_1]) \leq k, \text{ for every cycle } D \subseteq G[T_1]. \tag{19}$$

First suppose $G[T_1]$ is connected. By Proposition 24, there exists a path in $G[T_1]$ of length at least $2m$.

Claim 27. *Let $P = y_1 \ldots y_l$ be a path of maximum length in $G[T_1]$. If $y_i \in N(y_1)$ and $y_j \in N(y_l)$, then $i \leq j$.*

Proof. Suppose there exists $y_i \in N(y_1)$, $y_j \in N(y_l)$ such that $i > j$. With respect to this condition, choose y_i and y_j such that $i - j$ is minimum. If $i - j - 1 \leq \frac{1}{3}k$, set $D := y_1 \ldots y_jy_l \ldots y_iy_1$. By (14), $|D| \geq 2m - \frac{1}{3}k > k$, which contradicts (19). If $i - j - 1 > \frac{1}{3}k$, let h be maximum such that $y_h \in N(y_1)$ and set $D := y_1y_2 \ldots y_hy_1$. Since $i - j - 1 > \frac{1}{3}k$ and $i - j$ is minimum, we have $|D| \geq h \geq m + i - j - 1 > k$, which contradicts (19). $\square$

Let $P := y_1 \ldots y_l$ be a path of maximum length in $G[T_1]$ and with respect to this condition, choose P so that $j - i$ is minimum, where y_j is the smallest indexed neighbor of y_l and y_i the largest indexed neighbor of y_1. Note that by Claim 27, $j - i \geq 0$. By (19) we have,

$$N(y_1) \subseteq \{y_2, \ldots y_k\} \text{ and } N(y_l) \subseteq \{y_{l-k+1}, \ldots, y_{l-1}\}. \tag{20}$$

Set

$$A := \{y_1, \ldots, y_{i-1}\},\ B := \{y_i, \ldots, y_j\},\ C := \{y_{j+1}, \ldots, y_l\}.$$

Without loss of generality we may suppose $|A| \geq |C|$ and thus we have

$$m \leq \delta(G[T_1]) \leq |C| \leq |A| < k \tag{21}$$

and $|B| = j - i + 1 \leq l - 2m$.

Next we show that

$$\|A, C\| = 0. \tag{22}$$

Suppose $a < i \leq j < b$ and $y_ay_b \in E$. Choose $y_{a'} \in N(y_1)$ and $y_{b'} \in N(y_l)$ such that $a < a' \leq i \leq j \leq b' < b$ and both $a'-a$ and $b-b'$ are minimal. Now $D := y_1Py_ay_bPy_ly_{b'}Py_{a'}y_1$ is a cycle having the property that $N(y_1) \cup N(y_l) \subseteq V(D)$ and thus $|D| \geq |N(y_1) \cup N(y_l)| \geq 2m-1 > k$, contradicting (19).

Set $A' := \{y_h \in A : y_{h+1} \in N(v_1)\}$ and $C' := \{y_h \in C : y_{h-1} \in N(y_l)\}$. Note that $|A'| \geq m$ and $|C'| \geq m$. We claim that the vertices in $A' \cup C'$ are good in the sense that

$$\forall a \in A', N(a) \cap (T_1 \setminus (A \cup \{y_i\})) = \emptyset \text{ and } \forall c \in C', N(c) \cap (T_1 \setminus (C \cup \{y_j\})) = \emptyset. \tag{23}$$

Without loss of generality, suppose some $y_h \in A'$ has a neighbor $y' \in T_1 \setminus (A \cup \{y_i\})$. If $y' \notin V(P)$, then $y'y_h \ldots y_1y_{h+1} \ldots y_l$ is longer than P which is a contradiction. Otherwise, by

(22), $y' \in B$. However, $y_h \dots y_1 y_{h+1} \dots y_l$ is a path for which $j-i$ is smaller, contradicting the minimality of $j-i$.

Now suppose $G[T_1]$ is not connected. Since $\delta(G[T_1]) \geq m$ and $|T_1| < 3m$, $G[T_1]$ has exactly two components. Call these components A and C, then set $A' := A$ and $C' := C$. Without loss of generality, suppose $|A| \geq |C|$. Since $\delta(G[T_1]) \geq m$, we have $m+1 \leq |C|$ which implies $|A| < k$, by (14) and the fact that $|T_1| = 2k + \lfloor \alpha k \rfloor - |T_0|$. So regardless of whether $G[T_1]$ is connected or not, all of the following hold: (21), (22), (23), and

$$\forall a \in A', \overline{\|a, A\|} \leq |A| - m \quad \text{and} \quad \forall c \in C', \overline{\|c, C\|} \leq |C| - m. \tag{24}$$

For $Y \in \{A, C\}$, let $Y' = A'$ if $Y = A$ and let $Y' = C'$ if $Y = C$.

Claim 28. *For all $v \in V \setminus (A \cup C)$, there exists $Y \in \{A, C\}$ such that for all $y \in Y'$, $|(N(v) \cap N(y)) \cap Y| \geq 3$.*

Proof. For all $v \in V \setminus (A \cup C)$, we have

$$\|v, A \cup C\| \geq 2k - (|V| - (|A| + |C|)) = |A| + |C| - k. \tag{25}$$

Suppose there exists $v \in V \setminus (A \cup C)$ and $c \in C'$ such that $|(N(v) \cap N(c)) \cap C| \leq 2$. This implies that $\|v, C\| \leq |C| - m + 2$ by (24). So we have

$$\|v, A\| \geq |A| + |C| - k - (|C| - m + 2) = |A| + m - k - 2.$$

Let $a \in A'$, then by (14),

$$|(N(v) \cap N(a)) \cap A| \geq (|A| + m - k - 2) + m - |A| = 2m - k - 2 \geq \frac{1}{3}k \geq 3.$$

□

Claim 29. *There exist two disjoint square P^5's connecting edges of A to edges of C.*

Proof. Set $s := \lfloor \frac{|A|}{2} \rfloor$ and $t := \lfloor \frac{|C|}{2} \rfloor$. Choose nonadjacent vertices $x, x' \in S$ and $a_{2s}, c_1 \in N(x)$ with $a_{2s} \in A'$ and $c_1 \in C'$. Since a_{2s} and c_1 are nonadjacent they have at least $k+1$ common neighbors distinct from x, and these common neighbors are not in $A \cup C$. One of them v must also be adjacent to x. By Claim 28 there exists, without loss of generality, $a_{2s-1} \in A$ such that $a_{2s}, v \in N(a_{2s-1})$. Since $x \in S$, there exists $c_2 \in C$ such that $x, c_1 \in N(c_2)$. Thus $Q := a_{2s-1}a_{2s}vxc_1c_2$ is a square P^5 connecting $a_{2s-1}a_{2s}$ to c_1c_2. Similarly, we can choose $a_1, c_{2t} \in N(x')$ with $a_1 \in A' - a_{2s-1} - a_{2s}$ and $c_{2t} \in C' - c_1 - c_2$. Since a_1 and c_{2t} are nonadjacent, there exist k common neighbors of a_1 and c_{2t} that are distinct from x' and v. One of them v' is adjacent to x', and $v' \neq x$ by the choice of x, x'. Moreover, $v' \notin A \cup C$. So as above, we can choose $a_2 \in A$ and $c_{2t-1} \in C$ so that $Q' := c_{2t-1}c_{2t}\{v'x'\}a_1a_2$, $Q \cap Q' = \emptyset$ and Q' is a square P^5 connecting $c_{2t-1}c_{2t}$ to a_1a_2 (note that we cannot specify the order of v' and x'). □

Finally we claim that there exist paths

$$R := a_1a_2 \dots a_{2s-1}a_{2s} \subseteq G[A] \text{ and } R' := c_1c_2 \dots c_{2t-1}c_{2t} \subseteq G[C],$$

such that $|R| = 2s$ and $|R'| = 2t$. If $|A| = m$, then $A = A'$ and thus $G[A]$ is complete by (24). Otherwise $|A| \geq m + 1$ and thus by (14) we have

$$\frac{1}{3}k + 1 \leq \frac{1}{2}|A|. \tag{26}$$

By (22) and (26), we have

$$\delta(G[A]) \geq 2k - (|V| - (|A| + |C|)) = |A| + |C| - k \geq |A| + 2 - (\frac{k}{3} + 1) \geq \frac{1}{2}|A| + 2.$$

Thus for all $a, a', a'' \in A$,

$$G[A \setminus \{a, a', a''\}] \text{ is hamiltonian connected,}$$

since $\delta(G[A \setminus \{a, a', a''\}]) \geq \frac{1}{2}|A| - 1 > \frac{1}{2}(|A| - 3)$. If $|A| = 2s$, then we use the fact that $G[A \setminus \{a_1, a_{2s}\}]$ is hamiltonian connected to get R. If $|A| = 2s + 1$ we let $a' \in A \setminus \{a_1, a_2, a_{2s-1}, a_{2s}\}$, and we use the fact that $G[A \setminus \{a_1, a_{2s}, a'\}]$ is hamiltonian connected to get R. Since $|A| \geq |C|$, the same argument gives us R' in $G[C]$.

So by Claim 29, $D := RQR'Q'$ is an even cycle of length $2s + 2t + 4 \geq 2m + 2$ (note that $D \not\subseteq G[T_1]$). Recall that $V(D) \cap S \subseteq \{x, v, x', v'\}$ and set $S' := S \setminus D$. As in Case 1, let $P := \{p_1, \dots, p_s, p'_1, \dots, p'_t\}$ be a set of *ports*, where $p_i := \{a_{2i-1}, a_{2i}, a_{2i+1}, a_{2i+2}\}$ for $1 \leq i \leq s - 1$ and $p'_j := \{c_{2j-1}, c_{2j}, c_{2j+1}, c_{2j+2}\}$ for $1 \leq j \leq t - 1$. By Claim 26, there exist $x_1, \dots, x_{s-1}, x'_1, \dots, x'_{t-1}$ such that

$$a_1a_2x_1a_3a_4x_2 \dots x_{s-1}a_{2s-1}a_{2s}vxc_1c_2x'_1c_3c_4x'_2 \dots x'_{t-1}c_{2t-1}c_{2t}\{v'x'\}a_1a_2$$

is a square cycle of length at least $2s + 2t + 4 + s - 1 + t - 1 \geq 3m - 1 > 2k$. Thus by Theorem 4, G has a hamiltonian square cycle.

$\square$

5 Conclusion

We have established a concrete threshold $n_0 := 2 \times 10^8$ such that Pósa's Conjecture holds for all graphs of order at least n_0, using methods essentially from prior to 1996. It seems in retrospect, that we were blinded by the brilliance of the Regularity-Blow-up method, and missed that the crucial idea of [12] was just to divide the problem into extremal and non-extremal cases. However Pósa's Conjecture remains open. We suspect that our probabilistic methods cannot be used to obtain an improvement of more than a factor of 1000. On the other hand we believe that ordinary graph theoretic methods have not yet been exhausted.

We have also developed the method of special reservoirs, for removing regularity from certain arguments. We believe that this could be used on other problems. The paper [9] was

written with the goal of developing methods for a more general set of problems. In particular they used an *absorbing path* lemma which contributes to a much larger value of n_0. However other problems do not (yet) have an analog of Theorem 4, while the absorbing technique is quite adaptable. Here are some other possible candidates for applying these new techniques, the first of which was discussed in [9].

Conjecture 30 (Seymour [18])**.** *For all positive integers k, every graph G with $\delta(G) \geq \frac{k}{k+1}|G|$ contains the k^{th} power of a hamiltonian cycle.*

Komlós, Sárközy and Szemerédi [14, 15] used the Regularity and Blow-up Lemmas to prove that there exists a function $n(k)$ such that Seymour's Conjecture holds for all k and graphs of order at least $n(k)$.

Châu also used the Regularity and Blow-up Lemmas to prove the following Ore-type version of Pósa's Conjecture for graphs of large order.

Theorem 31 (Châu [1])**.** *Let G be a graph on n vertices such that $d(x) + d(y) \geq \frac{4}{3}n - \frac{1}{3}$ for all $xy \notin E(G)$.*

(a) *If $\delta(G) = \frac{1}{3}n + 2$ or $\delta(G) = \frac{1}{3}n + \frac{5}{3}$, then G contains a hamiltonian square path.*

(b) *If $\delta(G) > \frac{1}{3}n + 2$, then for sufficiently large n, G contains a hamiltonian square cycle.*

For a directed graph G, the *minimum semi-degree* of G, denoted $\delta^0(G)$, is the minimum of the minimum in-degree $\delta^-(G)$ and the minimum out-degree $\delta^+(G)$. An *oriented graph* is a directed graph with no 2-cycles. Keevash, Kühn, and Osthus proved the following oriented version of Dirac's theorem using the Regularity-Blow-up method (with a directed version of the Regularity Lemma).

Theorem 32 (Keevash, Kühn, Osthus [11])**.** *Let G be an oriented graph on n vertices. If $\delta^0(G) \geq \frac{3n-4}{8}$ and n is sufficiently large, then G contains a hamiltonian cycle.*

Finally Treglown conjectured the following oriented version of Pósa's conjecture.

Conjecture 33 (Treglown [20])**.** *Let G be an oriented graph on n vertices. If $\delta^0(G) \geq \frac{5n}{12}$, then G contains a the square of a hamiltonian cycle.*

Acknowledgements

We thank the referees for their thoughtful comments which improved the presentation of this paper.

References

[1] P. Châu, An Ore-type theorem on hamiltonian square cycles (2010), preprint.

[2] H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. Ann. Math. Statistics 23, (1952). 493–507.

[3] R. Diestel, *Graph Theory*, 4th Edition, Springer (2010).

[4] G. A. Dirac, Some theorems on abstract graphs, Proceedings of the London Mathematical Society 2 (1952) 68–81.

[5] P. Erdős, Problem 9, in: M. Fiedler (Ed.), *Theory of Graphs and Its Applications*, Czech. Acad. Sci. Publ., Prague, 1964, p. 159.

[6] G. Fan, H. A. Kierstead, The square of paths and cycles, Journal of Combinatorial Theory, Series B 63 (1995) 55–64.

[7] G. Fan, H.A. Kierstead, Hamiltonian square-paths, Journal of Combinatorial Theory, Series B 67 (1996) 167–182.

[8] G. Fan, H.A. Kierstead, Partitioning a graph into two square-cycles, Journal of Graph Theory 23 (1996) 241–256.

[9] I. Levitt, G. N. Sárközy and E. Szemerédi, How to avoid using the Regularity Lemma: Pósa's conjecture revisited, Discrete Mathematics 310 (2010) 630–641.

[10] S. Janson, T. Łuczak, A. Ruciński, Random Graphs, Wiley, New York, 2000.

[11] P. Keevash, D. Kühn and D. Osthus, An exact minimum degree condition for Hamilton cycles in oriented graphs, J. London Math. Soc. 79 (2009), 144–166.

[12] J. Komlós, G.N. Sárközy, E. Szemerédi, On the square of a hamiltonian cycle in dense graphs, Random Structures and Algorithms 9 (1996) 193–211.

[13] J. Komlós, G.N. Sárközy, E. Szemerédi, Blow-up Lemma, Combinatorica 17 (1) (1997) 109–123.

[14] J. Komlós, G.N. Sárközy, E. Szemerédi, On the Pósa-Seymour conjecture, Journal of Graph Theory 29 (1998) 167–176.

[15] J. Komlós, G. N. Sárközy, E. Szemerédi, Proof of the Seymour conjecture for large graphs, Annals of Combinatorics 2 (1998) 43–60.

[16] V. Rödl, A. Ruciński, E. Szemerédi, A Dirac-type theorem for 3-uniform hypergraphs, Combinatorics, Probability and Computing 15 (2006) 229–251.

[17] V. Rödl, A. Ruciński, E. Szemerédi, Dirac-type conditions for hamiltonian paths and cycles in 3-uniform hypergraphs, preprint.

[18] P. Seymour, Problem section, in: T.P. McDonough, V.C. Mavron (Eds.), Combinatorics: Proceedings of the British Combinatorial Conference 1973, Cambridge University Press, 1974, pp. 201–202.

[19] E. Szemerédi, Regular partitions of graphs, in: Colloques Internationaux C.N.R.S. No. 260 - Problèmes Combinatoires et Théorie des Graphes, Orsay, 1976, pp. 399–401.

[20] A. Treglown, A note on some embedding problems for oriented graphs, to appear in Journal of Graph Theory.